

# UNCLASSIFIED

AD NUMBER
AD860060
NEW LIMITATION CHANGE
TO Approved for public release, distribution unlimited
FROM Distribution authorized to U.S. Gov't. agencies and their contractors; Critical Technology; JUL 1969. Other requests shall be referred to Army Engineer Topographic Laboratory, Attn: ETL-GS [Geographic Information Systems], Fort Belvoir, VA 22060.
AUTHORITY
USAETL ltr, 1 Oct 1970

THIS PAGE IS UNCLASSIFIED

AD 860060

**SPECIAL INTERIM TECHNICAL REPORT  
THE BAYESIAN APPROACH TO IDENTIFICATION  
OF A REMOTELY SENSED ENVIRONMENT**

by  
**Robert Haralick**

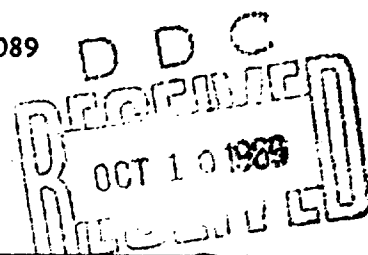
**CRES Technical Report No. 133-9  
July 1969**

Sponsored by  
**ADVANCED RESEARCH PROJECTS AGENCY, DEPARTMENT OF DEFENSE**  
STATEMENT #2 UNCLASSIFIED Work Order No. 1079

This document is subject to special export controls and each  
transmission to foreign government or foreign nationals may be  
made only with prior approval of

Monitored by  
**U.S. ARMY ENGINEER TOPOGRAPHIC LABORATORIES  
GEOGRAPHIC INFORMATION SYSTEMS BRANCH  
GEOGRAPHIC SYSTEM DIVISION  
FT. BELVOIR, VIRGINIA  
*attn: ETL.GS.***

Contract No. DAAK02-68-C-0089



CRES



THE UNIVERSITY OF KANSAS • CENTER FOR RESEARCH INC  
ENGINEERING SCIENCE DIVISION • LAWRENCE, KANSAS

69

AD \_\_\_\_\_

SPECIAL INTERIM TECHNICAL REPORT  
THE BAYESIAN APPROACH TO IDENTIFICATION  
OF A REMOTELY SENSED ENVIRONMENT

Robert Haralick

CRES Technical Report No. 133-9

July 1969

Sponsored by  
ADVANCED RESEARCH PROJECTS AGENCY, DEPARTMENT OF DEFENSE  
Work Order 1079

Monitored by  
U.S. ARMY ENGINEER TOPOGRAPHIC LABORATORIES  
GEOGRAPHIC INFORMATION SYSTEMS BRANCH  
GEOGRAPHIC SYSTEM DIVISION  
FT. BELVOIR, VIRGINIA

Contract No. DAAK02-68-C-0089

## ABSTRACT

The first part of this paper provides a brief tutorial introduction of the Bayesian Approach to identification of a remotely sensed environment. The second part describes the input data deck setup for the Fortran IV program which has been written to implement this approach. The third part describes file usage and subroutine organization. The fourth part provides a listing of the program with a simple sample data set.

## ACKNOWLEDGEMENT

I wish to acknowledge the assistance of Carl Smith who did most of the programming, and the University of Kansas Computation Center for providing some of the necessary computer time.

## TABLE OF CONTENTS

	<u>Page</u>
ABSTRACT	i
ACKNOWLEDGEMENT	ii
LIST OF ILLUSTRATIONS	iv
I THE BAYESIAN APPROACH TO IDENTIFICATION OF A REMOTELY SENSED ENVIRONMENT . . . . .	1
II INPUT DATA DECK . . . . .	7
III FILE USAGE AND SUBROUTINE ORGANIZATION . . . . .	15
IV EXAMPLE PROBLEM AND PROGRAM LISTING	49

## LIST OF ILLUSTRATIONS

		<u>Page</u>
Figure 1a	Flow Chart for Setting up Input Data Deck . . . .	13
Figure 1b	Flow Chart for Setting up Input Data Deck . . . .	14
Figure 2	Tape Usage . . . . .	16
Figure 3	Storage Requirements for the Tape or Disc Files .	17
Figure 4a	Data for Test 1 . . . . .	21
Figure 4b	Data for Test 2 . . . . .	21
Figure 5	Gain Matrix . . . . .	22
Figure 6	Input Data Deck for First Example Problem . . . .	25
Figure 7	Input Data Deck for Second Example Problem . . .	26

PART I

THE BAYESIAN APPROACH TO IDENTIFICATION  
OF A REMOTELY SENSED ENVIRONMENT



## I. THE BAYESIAN APPROACH TO IDENTIFICATION OF A REMOTELY SENSED ENVIRONMENT

Using remote sensors we can make measurements of an environment. The set of measurements made will be called the data set. Our job is to examine the data set in order to identify what the environment is made up of: our problem is how should we do it? In what follows we describe the Bayesian decision approach with a deterministic decision rule.

We assume that distinct boundaries enclose a limited environment, which is made up of small-area patches, one next to the other. The identification of the environment consists of identifying each small-area patch within one category of a given set of categories. We assume that such an identification is sensible and possible.

In order to make any identification we must have knowledge concerning which kind of measurements are typical measurements of the categories we wish to identify. This knowledge is succinctly contained in a classification, which is a mapping, associating with each measurement the category to which it is most typical - given a specific decision criterion. Therefore, if we are to identify measurements in a data set we must have a classification.

How do we obtain a classification? We perform an information gathering experiment. From the population of all environments, we sample one or a few in which it is possible to identify many small-area patches within each category of interest. The proportion of occurrence of each category in the sampled environment(s) does not have to be representative of the average probability of occurrence of each category in the entire environmental population. However, if we have no information regarding the average probability of occurrence of each category in the environmental population, then we would want to choose the sampled environment(s) so that the proportion of occurrence of each category in the sampled environment(s) is an unbiased estimate of the

average probability of occurrence of each category in the environmental population. In either case, the small-area patches within each of the sampled environment(s) do have to be representative of the categories with which they are identified.

With each of our sensors, we measure each small-area patch in the chosen environments. From photo-interpretation or field studies, the environments are examined first hand, and an identification of each small-area patch is made. The sequence of such identifications is called the "ground truth identification" or simply "ground truth". It is from the data set (the sequence of measurements) and the ground truth (the sequence of identifications) that we can find a Bayes classification.

At this point we must introduce some mathematical notation.

Let  $C = \{c_i\}_{i=1}^K$  be the set of  $K$  given categories;  $c_i$  is the symbol used

for the  $i^{\text{th}}$  category. We suppose, for convenience, that each sensor produces only one number for each measurement it makes of a small-area patch. We suppose further, that the  $j^{\text{th}}$  sensor must produce a number belonging to its range set  $L_j = \{l_{j1}, l_{j2}, \dots, l_{jN_j}\}$ . This supposition is fully in accord with reality, since the output of any sensor is always equivalent to a pointer-reading on a dial. Pointer-readings can never be discerned precisely, and are thus discerned approximately to third, or fourth, or, ...,  $N^{\text{th}}$  place accuracy.

Measurement space  $M$  is the set of all measurements which are possible to make with the set of  $S$  sensors.  $M$  is conveniently described as the cartesian product of the range sets;  $M = L_1 \times L_2 \times \dots \times L_S$ . This is the set of measurements which contain for elements, all the possible numbers produced by sensor one, combined with all the possible numbers produced by sensor two, ..., combined with all the possible numbers produced by sensor  $S$ . For convenience we number the measurements in  $M$ ;  $M = \{m_n\}_{n=1}^N$ , where  $N$  is the total number of elements in measurement space. Finally we must provide a goodness criterion; thus, we introduce a gain function  $g$ .  $g(c_i, c_j)$  is our economic gain if we identify a measure-

ment as belonging within the  $i^{\text{th}}$  category when that measurement was made of a small-area patch actually belonging within the  $j^{\text{th}}$  category.

We have already mentioned that a classification is a mapping or rule which associates with each measurement  $m_n$  in  $M$ , the category  $c_i$  to which it is most typical - according to some decision criterion. Our decision criterion is economic; "most typical to" translates to, "that association by which we, on the average, gain the most economically". Therefore, according to our decision criterion, we can judge each possible classification. That classification which enables us to gain the most, on the average, is the classification which is best; it is that classification which we wish to find.

Let us now examine how the average gain may be calculated. Let  $f$  be a classification mapping.  $f$  is a function whose domain is the set  $M$ , and whose range is the set  $C$ ;  $f: M \rightarrow C$ . For each element  $m_n \in M$  the function associates one and only one category  $c_i \in C$ . We define the characteristic function  $h_f$  for  $f$  as follows: for every  $m_n \in M$ ,  $c_i \in C$ ,

$$\left. \begin{aligned} h_f(c_i, m_n) &= 1 \text{ if and only if } f(m_n) = c_i \\ &0 \text{ otherwise} \end{aligned} \right\} .$$

In other words  $h_f(c_i, m_n)$  is 1 if and only if the classification  $f$  identifies the measurement  $m_n$  as belonging within the category  $c_i$ . The average gain  $A$  for the classification  $f$  is easily seen to be:

$$A(f) = \sum_{i=1}^K \sum_{k=1}^K \sum_{n=1}^N g(c_k, c_i) h_f(c_k, m_n) P(m_n | c_i) P(c_i)$$

where  $P(m_n | c_i)$  is the conditional probability that the measurement  $m_n$  will be made of a small-area patch given that the patch belongs within category  $c_i$ ,  $P(c_i)$  is the probability that any small-area patch of the

environments in the population belongs within category  $c_i$ , and  $g(c_k, c_i)$  is the amount gained if a patch which actually belongs within category  $c_i$  is identified within category  $c_k$ .

Of the four terms in the summation,  $g(c_k, c_i)$  is specified as part of the identification goodness criteria,  $h_f(c_k, m_n)$  is defined from the classification  $f$ ,  $P(m_n | c_i)$  will be determined from the data gathered in the experiment, and  $P(c_i)$  is an additional a priori probability which we will have to specify. Let us now examine in detail how the conditional probabilities are determined from the experimental data.

The data set is a sequence  $D$  of  $R$  measurements;

$$D = \langle m_{r_1}, m_{r_2}, \dots, m_{r_R} \rangle.$$

The ground truth corresponding to sequence  $D$  is a sequence  $T$  of  $R$  not necessarily different category identifications;  $T = \langle c_{r_1}, c_{r_2}, \dots, c_{r_R} \rangle$ .

Let  $\#$  be the counting measure.  $\#(D)$  is the number of elements in the sequence  $D$ ; thus,  $\#(D) = R$ . A sequence is really a function whose domain is the set of integers  $I$ . The data set  $D$  is then a function which associates with each integer, a measurement;  $D: I \rightarrow M$ . The ground truth  $T$  is also a function and it associates with each integer a category;  $T: I \rightarrow C$ .  $D(7)$ , for example, is then just the seventh element in the sequence  $D$ ;  $D(7) = m_{r_7}$ .  $D^{-1}(m)$  is the set of all integers  $i$  for which  $D(i) = m$ . The statistic  $\hat{P}(m_n | c_i)$  estimating  $P(m_n | c_i)$  is defined as

$$\hat{P}(m_n | c_i) = \left. \begin{aligned} & \frac{\#(D^{-1}(m_n) \cap T^{-1}(c_i))}{\#(T^{-1}(c_i))} \quad \text{when } \#(T^{-1}(c_i)) \neq 0 \\ & = 0 \quad \text{otherwise} \end{aligned} \right\}.$$

$\hat{P}(m_n | c_i)$  is the number of integers which are associated with the measurement  $m_n$  in the sequence D and with the category  $c_i$  in the sequence T, divided by the number of integers associated with the category  $c_i$  in sequence T. Stated simply,  $\hat{P}(m_n | c_i)$  is just the number of times the measurement  $m_n$  was made of a small-area patch belonging within category  $c_i$ , divided by the number of times a small-area patch belonged within the category  $c_i$ .

The a priori probabilities  $P(c_i)$  can either be estimated from the sampled data set (if this data set is representative of the population) or from our foreknowledge of the population of environments. If we can assume that the few environments we have chosen to sample for our experiment are representative of the population, then

$$\hat{P}(c_i) = \frac{\#(T^{-1}(c_i))}{R}$$

is a reasonable estimate. If we cannot make such an assumption and we believe that a small-area patch is just as likely to belong within one category as within another, then  $\hat{P}(c_i) = 1/K$  is a reasonable estimate.

From the estimates  $\hat{P}(m_n | c_i)$  and  $\hat{P}(c_i)$  we may estimate the average gain  $\hat{A}$  for any classification  $f$ . As before let  $h_f$  be the characteristic function for  $f$ .

$$h_f(c_k, m_n) = \begin{cases} 1 & \text{if and only if } f(m_n) = c_k \\ 0 & \text{otherwise} \end{cases}$$

$$\hat{A}(f) = \sum_{i=1}^K \sum_{k=1}^K \sum_{n=1}^N g(c_k, c_i) h_f(c_k, m_n) \hat{P}(m_n | c_i) \hat{P}(c_i).$$

We seek the Bayes classification  $f^*$  which maximizes  $\hat{A}$ .  $f^*$  is easily defined. For each measurement  $m_n$  and for any classification  $f$ ,

there will be one and only one category  $c_j$  such that  $h_f(c_j, m_n) = 1$ . Consider the amount  $\hat{a}(c_j, m_n)$  gained due to the identification of measurement  $m_n$  as belonging within category  $c_j$ .

$$\hat{a}(c_j, m_n) = \sum_{i=1}^K g(c_j, c_i) \hat{P}(m_n | c_i) \hat{P}(c_i)$$

The maximum  $\hat{A}(f)$  is certainly achieved if for each measurement  $m_n$ ,  $f(m_n) = c_j$  where  $c_j$  maximizes  $\hat{a}(c_j, m_n)$ . Therefore we just have to compute  $\hat{a}(c_j, m_n)$  for  $j = 1, 2, \dots, K$  to determine which category,  $c_j$ , maximizes it. Then we define  $f^*(m_n) = c_j$ .

In this manner we can define how to best identify each measurement which actually occurred in the data sequence  $D$ . However, there may be many measurements in measurement space  $M$  which did not occur in the data sequence. How should these measurements be identified in the classification? Since we have no data or statistics for these measurements it seems that we have no way to deal with them! Here we must draw upon our knowledge of the structure of reality. We know that in any environment if a measurement  $m$  is made of a small-area patch belonging within category  $c_i$ , then it is likely to make measurements  $m + \delta$  for other small-area patches which also belong within category  $c_i$ . If a measurement  $m$  is typical of category  $c_i$ , then for small  $\delta$ ,  $m + \delta$  is also typical of category  $c_i$ . Similar or close measurements are usually associated with similar or the same categories. Thus in the classification we can identify a measurement  $m$ , which did not occur in the data sequence, with the category associated with  $m'$ , its nearest neighbor.

The part of the classification  $f^*$  which was defined by means of the statistics generated by the experiment is called a Bayes Classification and hence the name "Bayesian approach." The part of the classification which is not Bayesian is said to be defined by a nearest neighbor search.

**Acknowledgement:** This work was supported by Project Themis (USAETL Contract DAAK02-68-C-0089, ARPA order No. 1079).

## PART II

### INPUT DATA DECK

## II. INPUT DATA DECK

The data for this program are received as a sequence of measurements of small-area patches or objects with each measurement made by a sensor or set of sensors. A measurement may be, for example, the average backscatter power return from a small-area patch at incidence angles 5°, 10°, 15°, 20°, 30°, 40°, 50°, and 60°. In this case, each measurement has 8 components. The patches themselves are examined and identified as belonging within one of several given categories. The sequence of such identifications is called the "ground truth identification" or simply the "ground truth." The Bayes program can determine a Bayes classification of measurement space, based on the data and the ground truth for the data. Once a classification is determined, the Bayes program can identify each measurement within a sequence of measurements. This identification is done by a nearest-neighbor search.

The input deck is organized into four sections: title, program options, parameter cards, and format and data.

### I. Title

- A) The title section consists of a single card specifying the name of the data. The title may begin in column one and continue through column eighty.

### II. Program Options

- A) The program option section consists of two cards, the first card specifying all the input options and the second card specifying all the output options.
- B) Each option is a six-character abbreviation or code.
- C) The options start in column 16, are separated only by commas (no embedding blanks) and may appear in any order.
- D) Only input options may appear on the input card and only output options may appear on the output card.
  - 1) The input options are: PHOPTS, CORPTS, FLTING, FATERN, DIAGON, HLFNHF, ABSQNT.
  - 2) The output options are: ALTICH, STDPNT, DPUNCH, PHOUT1, PHOUT2, TERMINL.



### III. Parameters

- A) The parameter section consists of cards, the number of which varies with the options chosen.
- B) There are seven basic types of information which can possibly appear in the parameter section: gain matrix, dimensionality of measurements, number of measurements in the data set, number of categories in the classification, display size, number of levels to which the measurements will be quantized, and means of estimating the a priori probability distribution.

### IV. Format and Data

- A) There are two ways to organize the format and data: photographic form and corresponding-point form. Depending on the options chosen, the ground truth identification and its format may or may not be present. One and only one of the two forms must be specified: otherwise, an error message and termination of the job will result.
  - 1) In the photographic form the data are organized as follows:
    - a) format for identification (if any)
    - b) identification (if any) for measurement one, measurement two, ..., measurement N.
    - c) format for measurements
    - d) component one, measurement one; component one, measurement two; ...; component one, measurement N; component two, measurement one; component two, measurement two; ...; component two, measurement N; ... component M, measurement one; component M, measurement two; ...; component M, measurement N.
  - 2) In the corresponding-point form the data are organized as follows:
    - a) format for identification (if any) and measurements

- b) identification (if any) for measurement one; component one, measurement one; component two, measurement one;...; component M, measurement one; identification (if any) for measurement two; component one, measurement two; component two, measurement two;... component M, measurement two:... identification (if any) for measurement N; component one, measurement N; component two, measurement N;... component M, measurement N.

We now describe the options.

#### I. Input Options

- A) PHOPTS -- is the abbreviation for photographic form, and is used when the data are in that form.
- B) CORPTS -- is the abbreviation for corresponding point form, and is used when the data are in that form.
- C) FLTNG -- is the abbreviation for floating point, and is used when the data are punched on cards in floating-point form. It is not used when the data are punched on cards in integer form.
- D) PATTERN -- is the abbreviation for pattern classification by Bayes' strategy. Use of this option will output a probability matrix where element  $(i, j)$  is the conditional probability that a measurement which was identified as within the  $i^{\text{th}}$  ground truth category is identified in the classification as within the  $j^{\text{th}}$  category. A punched deck of the compacted quantized measurements with their identifications in the Bayes' classification will also be produced.
- E) DIAGON -- is the abbreviation for diagonal gain matrix with ones on the diagonal. Specification of DIAGON will internally generate an identity matrix for the gain matrix. This relieves the user of the need to supply the appropriate cards in the parameter section.

- F) HLFNHF -- is the abbreviation for half and half. Specification of HLFNHF will divide the data into halves: even and odd points. The first, third, . . . , data points are used to construct a Bayes' classification, and the second, fourth, . . . , data points are identified on the basis of the classification.
- G) ABSQNT -- is the abbreviation for absolute maximum quantization. Specification of ABSQNT will quantize the measurements by determining the minimum and maximum values occurring among all the components, normalizing each component by subtracting this minimum, dividing by this maximum minus this minimum, and multiplying by the number of quantized levels desired. If ABSQNT is not specified, the program finds the maximum and minimum for each component, and normalizes component by component.

## II. Output Options

- A) ALTPCH -- is the abbreviation for alternate punch. Specification of ALTPCH produces a punched quantized data deck in the alternate form. If the data are in photographic form, the punched deck will be in corresponding-point form. If the data are in corresponding-point, the punched deck will be in photographic-point form.
- B) STDPNT -- is the abbreviation for standard identification print out. Specification of STDPNT will print out the ground truth identification.
- C) DPUNCH -- is the abbreviation for punched deck. Specification of DPUNCH will produce a punched quantized data deck in the same form as the input data.
- D) PHOUT1 -- is the abbreviation for photograph output. Specification of PHOUT1 will identify data according to a given classification. The classification can be internally generated by the Bayes routine or it can be externally supplied in the data deck. Identification is done by a

table look-up procedure. If the quantized measurement cannot be found in the classification, a nearest-neighbor search is initiated, and the measurement is identified with the same identification as the closest measurement to it in the classification.

If the input includes ground truth identification for a data set and the data are further identified relative to a classification (by specification of PHOUT1), then a contingency table of the ground truth identification versus the classification identification will be printed out. The  $(i,j)^{th}$  element of the table is the number of measurements which were identified in the  $i^{th}$  ground truth category and classified in the  $j^{th}$  ground truth category.

- E) PHOUTD -- is the abbreviation for photograph output deck. Specification of PHOUTD will produce a punched deck of the identification which the photograph output routine determined. PHOUTD can only be specified if PHOUT1 is specified.
- F) TERMNL is the abbreviation for remote terminal format. Specification of TERMNL will format all printed output from the Bayes program so that each line has no more than seventy columns. If TERMNL is not specified, each line is printed with one hundred thirty columns.

The generality of the program requires that the input parameters be quite variable, depending on the type of data to be processed and the options desired. For example, the PATTERN option may or may not require the gain matrix as input depending on whether or not DIAGON was also specified. To facilitate setting up the input data deck, a flow chart is illustrated in Figure 1a and 1b.

The flow chart contains two symbols. The first symbol is the elongated circle (rectangle with rounded edges), which asks the question printed in the circle and requires a "true" or "false" answer. Depending upon the user's answer, one branch is chosen from the bottom of the circle.

The second symbol is the rectangle, which usually signals the addition of a single data card. Data cards are added when rectangles are encountered in the flow chart. Two exceptions to the "one-rectangle, one-data-card" rule are the gain matrix and the data itself. If the gain matrix is required (PATTERN is specified and DIAGON is not specified), the user must supply all the elements for the matrix. This will take up more than one card if there are more than two ground truth identification groups. Similarly, if there are more than eight measurements in the data set, the data deck will have more than one card.

The final control card is the STOP card and is the last card in the input deck. The four-character word STOP is punched in columns one through four. The program is designed to handle more than one data set per run, and the user may place behind the first (or second, or third, etc.) data set the NAME of the next data set, the proper /INPUT and /OUTPUT cards, and the other necessary parameter cards as specified by the flow chart. The STOP card is placed after the final set of data, and informs the system to terminate the job. Excluding the STOP card or misplacing any control cards will cause a read error and termination of the job.

# FLOW CHART FOR SETTING UP INPUT DATA DECK

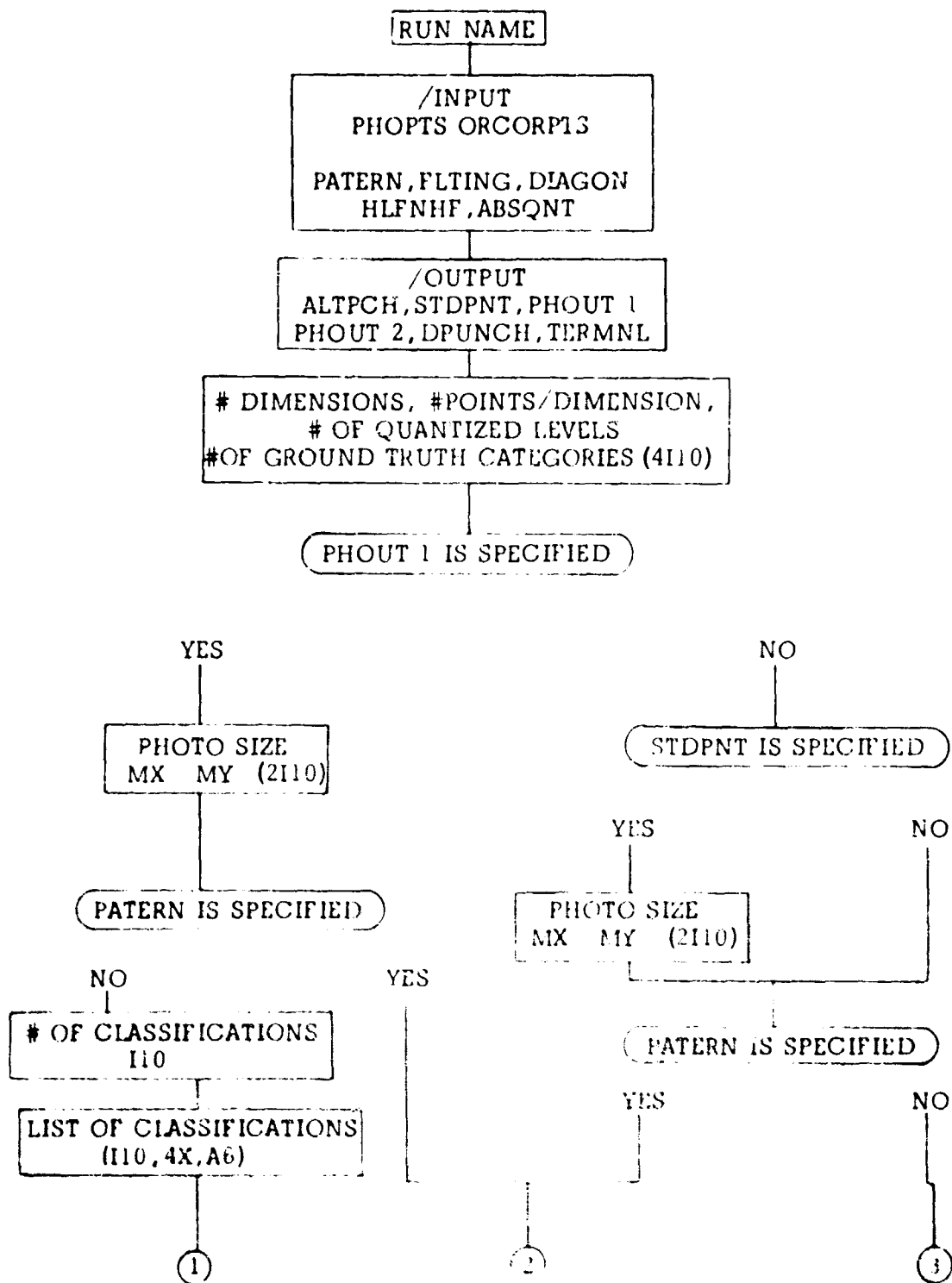


Figure 1a.

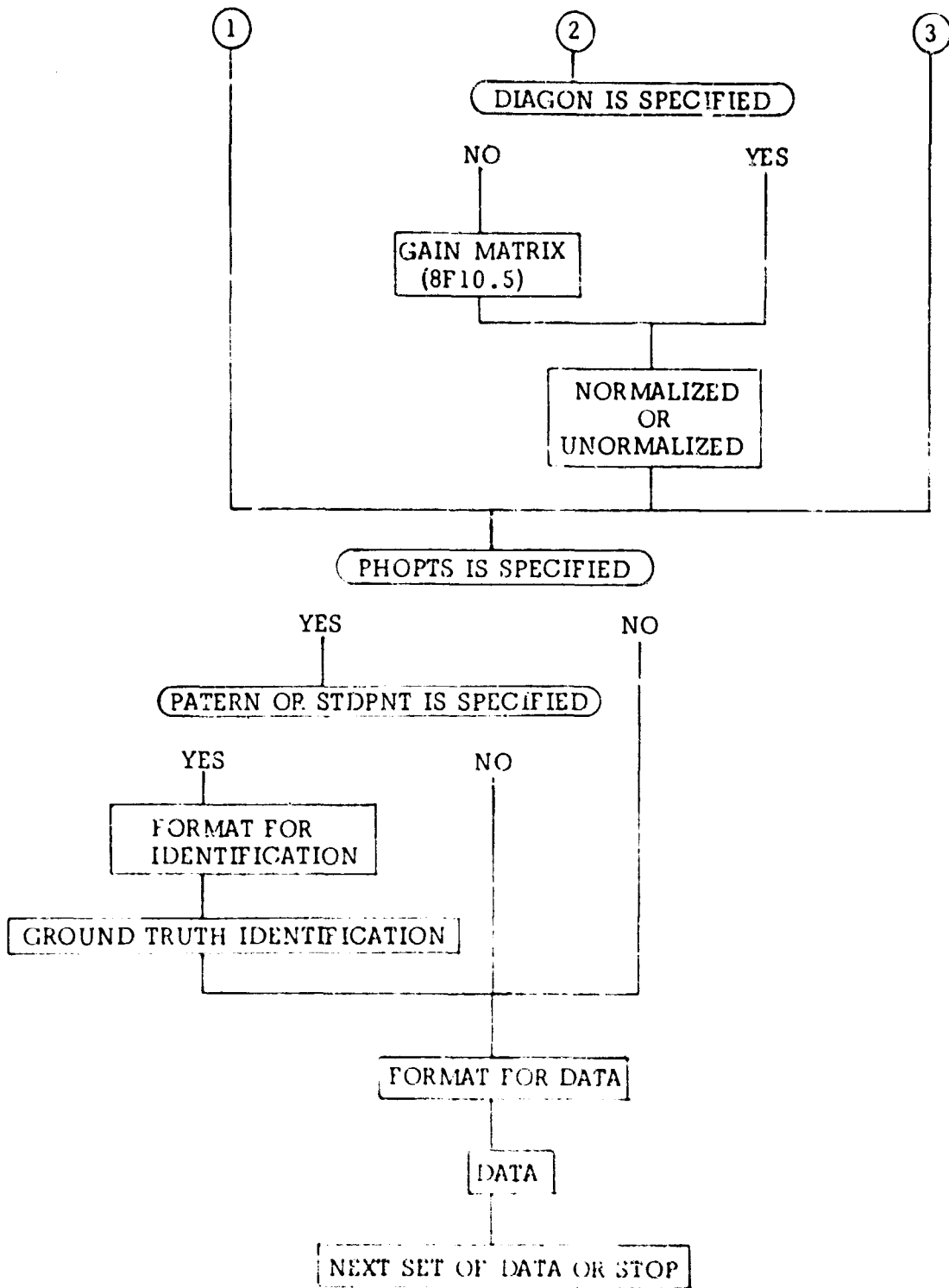


Figure 1b.

## PART III

### FILE USAGE AND SUBROUTINE ORGANIZATION



### III. FILE USAGE AND SUBROUTINE ORGANIZATION

Blank common storage carries all problem parameters and user options (CORPTS, PHOPTS, etc.), as well as providing a 24,000-word scratch area. Many routines in different links require such parameters as the number of dimensions of the current problem, the number of points being processed, and the number of ground truth categories.

All other communications between links are handled by tape, disc, or drum files. The program requires nine files -- 01, 02, 03, 04, 09, 10, 11, 20, 21 -- as well as the input (05), output (06), and punch (43) files normally used in FORTRAN. Figure 2 describes file usage, and Figure 3 illustrates how much storage is needed on each of the files.

The Bayes program, as mentioned before, requires 36,000 words of storage in the computer core, of which 25,000 may be shared during loading. The program has been observed to process 350 sixteen-dimensional data points in less than ten minutes' processor time.

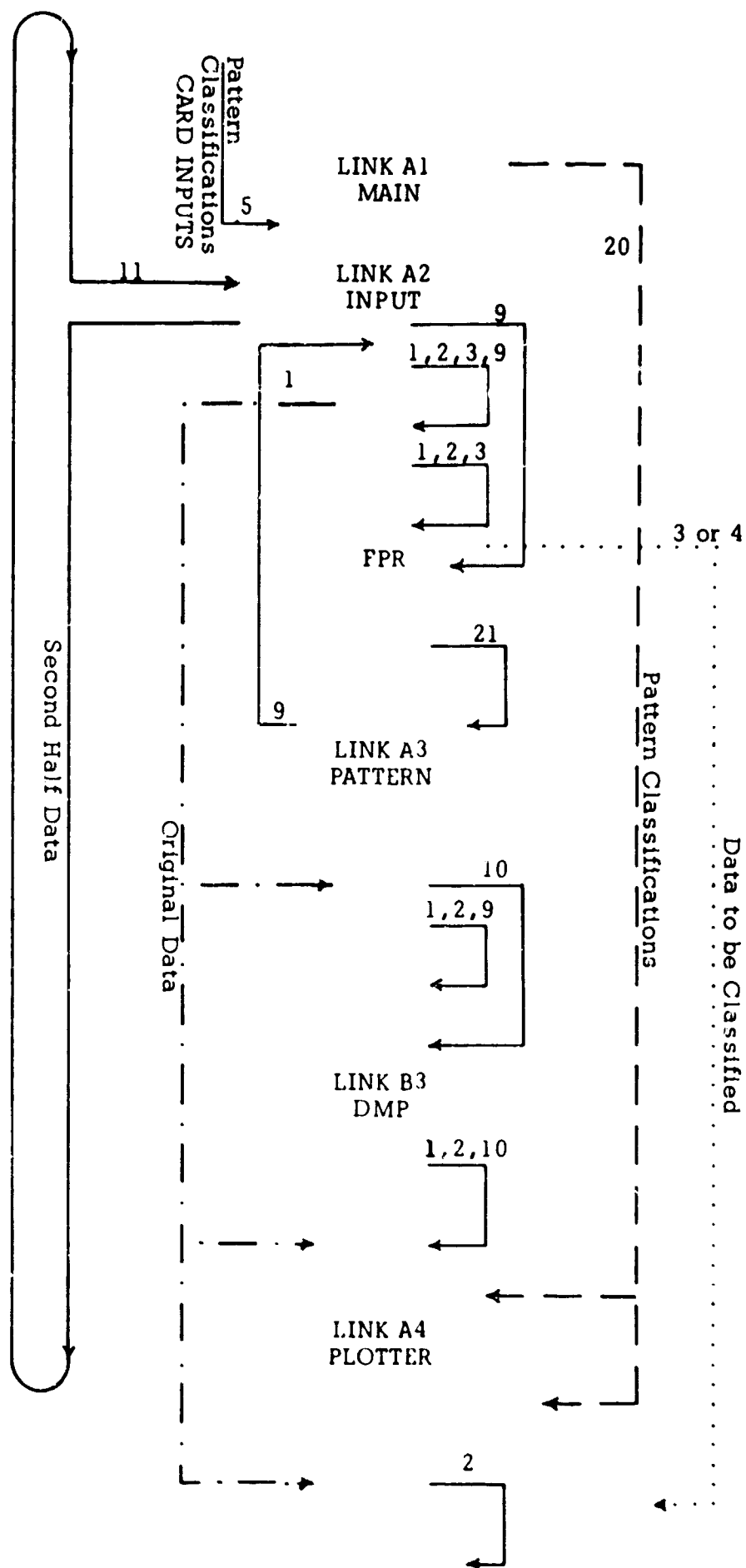


Figure 2. Tape Usage

FILE	NUMBER OF WORDS NEEDED
01	Total points
02	Total points
03	(Total Points) * # dimensions
04	
09	(Total Points) * # dimensions
10	2*(total points)
11	1/2*(total point)
20	2*(# unique n-tuples in data)
21	(Total Points) * # dimensions

Figure 3. Storage Requirements for the Tape or Disc Files

The program organization is briefly described below. Listed first are the mainline and CHNXT, the system supervisor, followed by the six-character link names used during overlay processing. Under each link name are listed the routines contained in the link. Link names containing the alphabetic character "A" refer to links essential for proper data processing; link names containing the alphabetic character "B" refer to links which output intermediate calculations, but which do not contribute to the overall program results.

#### LINKB2

.....(mainline) -- reserves all common storage, provides entry point to program, and contains comment cards stating program parameters and input deck setup.

CHNXT -- a small resident program used to control the entry and exits of the different links. This routine must comply with the overlay rules of the operating system in use.

#### LINKA1

MAIN -- parameter input and selection is accomplished in this routine. Also, if the photo classifications and the gain matrix must be read in, it is done in MAIN.

FORM -- selects and places in common all formats affected by terminal or non-terminal use.

#### LINKA2

INPUT -- performs data input, saves every even point for later processing if desired, and changes data to alternate form if called for.

FPR-FPR1 -- searches data for the maximum and minimum points so that proper quantization and shifting may be done in INPUT.

DEF -- defines the single-character symbols to be used in the classification.

TRANE -- prints a cross-reference between the single-character symbols used in the program and the original symbols.

CHANGE -- an assembly language routine which creates a suitable output format for the data.

#### LINKA3

PATERN -- Bayes program

OUTP -- prints out results of PATERN in eye-appealing format.

DECSON -- selects and assigns to each n-tuple the proper classification according to Bayes theory.

#### LINKB3

DMP -- used for debugging; prints out n-tuples vs. categories and n-tuples vs. classification.

#### LINKA4

PLOTTER -- classifies and plots input data according to n-tuple classifications.

SEARCH -- searches a list of n-tuples to find the list element which is closest in distance to another n-tuple.

IDIST -- calculates n-dimensional space distances.

## **PART IV**

### **EXAMPLE PROBLEM AND PROGRAM LISTING**

#### IV. EXAMPLE PROBLEM

Suppose the problem is to input a set of data in photographic form, punch out the data in the alternate form, print the data identification, and then classify a new set of data. Let there be nine points in the horizontal direction and ten points in the vertical direction for the first data set. The data appear in Figure 4a. For the first part of the problem, the program must produce an alternate data deck, a print-out of the identification, and a Bayes' classification based on the first set of input data.

The first card is the title card. The next two cards specify the input and output options needed. The options start in column 16 and are separated by commas. The data are in photographic form, so the input card is:

```
/INPUT      PHOPTS,PATERN
```

The output card is:

```
/OUTPUT      ALTPCH,STDPNT
```

The parameter card follows. From the flow chart we see that the fourth card must specify the number of dimensions per measurement (number of photographs), total number of measurements, number of quantized levels, and number of ground truth categories. In our example the number of photographs is two and the number of measurements is ninety. We wish to quantize the data to ten levels and there are two ground truth categories. The fourth card thus appears as:

```
2          90          10          2
```

Since PHOUT1 is not specified and STDPNT is, the fourth card must specify the photograph size, which is nine points horizontal by ten vertical. The fifth card appears as:

```
9          10
```

111151111  
111151111  
111151111  
111151111  
555555555  
555555555  
111151111  
111151111  
111151111  
111151111  
111151111

Photo 1 for Test 1

222262222  
222262222  
222262222  
222262222  
333333333  
333333333  
222262222  
222262222  
222262222  
222262222

Photo 2 for Test 1

Figure 4a. Data for Test 1

51115  
15151  
11511  
15151  
51115

Photo 1 for Test 2

32223  
26262  
22322  
23232  
62226

Photo 2 for Test 2

Figure 4b. Data for Test 2



Since PATTERN is specified and DIAGON is not, the user must supply the gain matrix. Suppose we choose a gain matrix where we gain ten for a correct decision and lose five for an incorrect decision, as illustrated in Figure 5.

$$\begin{matrix} 10 & -5 \\ -5 & 10 \end{matrix}$$

Figure 5 . Gain Matrix

The sixth card, specifying the above matrix, is

10.            -5.            -5.            10.

We must now indicate how the a priori probability distribution is estimated. We choose to suppose that each identification group has equal probability. Therefore NORMALIZED is specified on the next card,

NORMALIZED

The data are in photographic form, so the format for the ground truth identification must come next. In our example this would be:

(9A1).

After the identification format, the identification itself comes:

AAAAA  
AAAAA  
AAAAA  
AAAAA  
BBBBB  
BBBBB  
AAAAA  
AAAAA  
AAABAAA  
AAAAA

Finally we reach the format for the data, the data itself, and the STOP card.  
In our example these appear as:

```
(911)
111151111
111151111
111151111
111151111
555555555
555555555
111151111
111151111
111151111
111151111
111151111
222262222
222262222
222262222
222262222
333333333
333333333
222262222
222262222
222262222
222262222
STOP
```

The input data deck is illustrated in Figure 5.

At this point we must run a job with the input deck as shown in Figure 6. The job produces a punched deck of the Bayes classification. We must now identify a new set of data which is illustrated in Figure 4b. This is done by a separate job. The first card is, as usual, the title card. The second and third cards are the input and output option cards. The new set of data is in photographic form, and we wish to have a print-out of the identification for it, based on the classification of the previous job. The next cards thus appear:

```
/INPUT      PHOPTS
/OUTPUT     FHOUT1
```

There are two dimensions, twenty-five measurements, ten quantized levels, and two identification groups. The fourth card is:

2            25            10            2

Since PHOUT1 is specified, the next card must indicate the photographic size which is, in our example, five points horizontal by five vertical. The next card is thus:

5            5

The number of quantized measurements in the classification and the Bayes classification itself come next. These cards were obtained from the output of the previous job.

For our example, they are:

3            65.02            35.02            21.01

All that now remains is the format for the data, the data itself, and the STOP card.

(511)  
51115  
15151  
11511  
15151  
51115  
32223  
26262  
22322  
23232  
62226  
STOP

The input data deck is illustrated in Figure 7.



STOP  
62225  
62224  
62222  
62260  
62223  
51115  
15151  
11511  
15151

21 1  
35 2  
65 2  
3 5  
5 25 10  
2

PHOUT1  
PHOUT2

OUTPUT  
INPUT  
TEST DECK 2

00000

Figure 7. Input Data Deck for Second Example Problem



Line	Code	Statement	Column	Row
1		SUBROUTINE CMNXT		1
2		DIMENSION NAME(3)		2
3		COMMON MPD1,FORMAT1(12)		3
4		COMMON F		4
5		COMMON FMT1(15),FMT2(3),FMT9(5),FMT11(6),FMT12(7),FMT13(3),FMT14(2		5
6		1),FMT16(2),FMT22(3),FMT56(6),FMT97(4),FMT98(8)		6
7		COMMON L8,L9,L10,L11,L12,L13,L14		7
8		COMMON N,M,LEVEL,KZ1,KCOUNT(36),ITAPE,ANAY(625),LT,L1,L2,L3,L4,L5,		8
9		L6,KK,MX,MY,K,L7,MA,MB,MC,ZZ(24000)		9
10		LOGICAL LT,L1,L2,L3,L4,L5,L6,L7		10
11		LOGICAL L8,L9,L10,L11,L12,L13		11
12		INTEGER F		12
13		DATA NN/6NSTOP /		13
14	C		...	14
15	C	K POINTS TO LINK IN USE		15
16	C		...	16
17		1 K=K+1		17
18		GO TO (2,3,4,5,6,7), K		18
19		2 READ (F,9) NAME		19
20		IF (NAME(1).EQ.NN) STOP		20
21	C	CALL MAIN		21
22		CALL LINK (6MLINKA1)		22
23	C	CALL INPUT		23
24		3 CALL LINK (6MLINKA2)		24
25		4 IF (L6) GO TO 6		25
26		CALL LINK (6MLINKA3)		26
27		5 IF (L6) GO TO 6		27
28		CALL LINK (6MLINKA3)		28
29		6 K=5		29
30		IF (L7) GO TO 7		30
31		CALL LINK (6MLINKA4)		31
32		7 IF (L12) GO TO 8		32
33		K=8		33
34		GO TO 1		34
35	C		...	35
36	C	IF DATA IS RUN IN HALFS SET UP PARAMETERS		36
37	C		...	37
38		8 L12=.FALSE.		38
39		L13=.TRUE.		39
40		L5=.FALSE.		40
41		L6=.TRUE.		41
42		L7=.FALSE.		42
43		M=MPD1/2		43
44		K=1		44
45		WRITE(6,100) M		45
46		100 FORMAT(/'/34M SECOND HALF. NUMBER OF POINTS IS ,110)		46
47		GO TO 1		47
48		ENTRY MNAME		48
49	C		...	49
50	C	PRINT ID NAME		50
51	C		...	51
52		WRITE (6,FMT22) NAME		52
53		RETURN		53
54	C		...	54
55	C		...	55
56	C		...	56
57		9 FORMAT (346)		57
58		END		58

28

29



ՀԱՅԿԱՆ ԷՄԹԵՆՆԻԿ ՊԵՐՏՈՐ

39

1	SUBROUTINE FFRM(L)			1
2	COMMON MP01,F0RMT1(12)			2
3	COMMON F			3
4	COMMON FMT1(19),FMT2(3),FMT0(9),FMT11(6),FMT12(7),FMT13(3),FMT14(2)			4
5	1,FMT16(2),FMT22(3),FMT56(6),FMT57(4),FMT58(8)			5
6	DIMENSION O(64), T(64)			6
7	LOGICAL L			7
8	INTEGER F			8
9	DATA T/6M//39X,6M,11MLA,6MBLE TA,6MBLE //6M 16X,2,6MMPLOT,6M 3V			9
10	1B0,6MLS VS,6MINPUT,6MSYNROL,6MS //2,6M6X,A1,6M12X,A6,6M)			10
11	2,6M,6M(26X,6M4A1,12X,6M,A6),6M(27X,1,6MONTA1,6MNING R,6M			11
12	3EIONS,6M //),6M(24X,2,6MMPHOT,6MO CLAS,6MSIFICA,6MNTION /,6M//			12
13	4,6M//15X,6M,29MTR,6Maining,6M REGIO,6MN CLAS,6MSIFIED,6M AS //)			13
14	5,6M(5X,6(1,6M5X,12,6M//),6M(1M,2,6M,A1),6M(3X,A1,6M),6M			14
15	6(1M,2,6M5X,3A6,6M//),6M(24X,2,6M2MTH,6MPROB6M,6MILITY,6MMAT			15
16	7R1X,6M),6M(6X,12,6M,6(2X,6MF4,2),6M,6M(/////6M//21X,			16
17	8,6M17MOT,6MAL PRO,6MBABLI,6MNTY //2,6M3X,1E1,6M2,4) /			17
18	DATA O/6M//60X,6M,11MLA,6MBLE TA,6MBLE //6M51X,31,6MMPLOT,6MSYN			18
19	1BOL,6MS VS,6MINPUT,6MSYNROL,6MS //56,6M6X,A1,1,6M2X,A6,6M			19
20	2,6M,6M(56X,6M4A1,12X,6M,A6),6M(57X,1,6MONTA1,6MNING R,6M			20
21	3EIONS,6M //),6M(54X,2,6MMPHOT,6MO CLAS,6MSIFICA,6MNTION /,6M//			21
22	4,6M//58X,6M,29MTR,6Maining,6M REGIO,6MN CLAS,6MSIFIED,6M AS //)			22
23	5,6M(4X,6(1,6M13X,12,6M//),6M(1M,1,6M,A1),6M(2X,A,6M1),6M			23
24	6(1M,5,6M5X,3A6,6M//),6M(54X,2,6M2MTH,6MPROB6M,6MILITY,6MMAT			24
25	7R1X,6M),6M(5X,12,6M,6(3X,6ME12,4),6M),6M(/////6M//51X,			25
26	8,6M17MOT,6MAL PRO,6MBABLI,6MNTY //6M/53X,1,6ME12,4)///			26
27		..		27
28	SETUP FORMATS FOR USE IN THE PROGRAM			28
29		..		29
30	IF (L) GO TO 2			30
31	DO 1 I=1,64			31
32	1 FMT1(I)=F(I)			32
33	RETURN			33
34	2 DO 3 I=1,64			34
35	3 FMT1(I)=O(I)			35
36	RETURN			36
37	END			37

31

### CONFIDENTIALITY REPORT

32



00370 01 05-25-60 21.006

COMPACT EXECUTION REPORT

```

10 REMIND 0
11 TO 20 J01.1
12 ACCUMULATE (DATA(1),0.0) + (DATA(1),0.0)
13 TO 20 J01.1
14 READ (0) (DATA(1),0.0)
15 TO 20 J01.1
16 DATA(1) = (DATA(1) - 1.0) / (DATA(1) - 1.0)
17 CONTINUE
18 TYPE 3 MAT QUANTIZED DATA
19 WRITE (3) (DATA(1),0.0)
20 CONTINUE
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

00379 01 05-25-74 21.000

PHASE EXECUTION REPORT

```

C      USE ALL DATA FROM FIRST HALF
WRITE (11) TRAINING(DATASET), (I,1,N)
GO TO 36
36  WRITE (76) COUNT TRAINING(DATASET), (I,1,N)
IF (COUNT) GO TO 37
C      DATA IS IN INTERIOR FORM
DO 37 (I,1,N)
IF (DATA(I),LT,LSCALE) LSCALE=DATA(I)
37  IF (DATA(I),GT,LAMPS(I)) LAMPS(I)=DATA(I)
38  WRITE (9) (DATA(I),1,N)
IF (COUNT) GO TO 40
JUMPOVER
DO 38 (I,1,N)
JUMPOVER
381  TRAINING(JUMPOVER)=DATA(I)
CONTINUE
40  CONTINUE
401  CONTINUE
C      LA IS TRUE IF DATA IS IN INTERIOR FORM
IF (LA) GO TO 398
CALL FORM (LSCALE,LAMPS)
GO TO 39
398  IF (COUNT) GO TO 37
C      DATA IS TO ABSOLUTELY QUANTIZED
LARGE=LAMPS(I)
LSCALE=LSCALE
DO 398 (I,1,N)
IF (LARGE,LT,LAMPS(I)) LARGE=LAMPS(I)
IF (LSCALE,GT,LSCALE) LSCALE=LSCALE
700  CONTINUE
DO 398 (I,1,N)
LARGE=LARGE
LSCALE=LSCALE
701  LSCALE=LSCALE
C      ***
C      QUANTIZE DATA AND BUILD TRAINING REGION ARRAY WITH COUNT = 0
C      ***
39  REWIND 0
DO 40 (I,1,N)
40  ACCUMULATE (LSCALE,LSCALE,DATA,LARGE,LSCALE)
DO 40 (I,1,N)
READ (9) (DATA(I),1,N)
DO 41 (I,1,N)
DATA(I)=DATA(I)-LSCALE+LARGE
41  CONTINUE
WRITE (13) (DATA(I),1,N)
42  CONTINUE
REWIND 2
STORE (DATA(I))
COUNT=COUNT+1
DO 42 (I,1,N)
43  IF (COUNT,GT,COUNT) GO TO 44
COUNT=COUNT
IF (COUNT) GO TO 43
STORE (DATA(I))
COUNT=COUNT+1
DO 43 (I,1,N)
44  COUNT=COUNT+1
45  REWIND 3
IF (COUNT) GO TO 42
DO 45 (I,1,N)
DO 45 (I,1,N)
REWIND 3
45  CONTINUE
C      ***
C      STORE TRAINING REGION IN TAPE 3
C      ***
DO 46 (I,1,N)
READ (13) (DATA(I),1,N)
IF (COUNT,GT,COUNT) GO TO 46
DATA(I)=DATA(I)
46  CONTINUE
C      ***
C      STORE DATA ON TAPE 2
C      ***
47  WRITE (12) (DATA(I),1,N)
C      ***
C      QUANTIZED DATA IN INTERIOR LEVEL FORM 3 THROUGH 5
C      PLACE RESULTS ON TAPE ONE DIMENSION BY DIMENSION
C      ***
48  REWIND 3
REWIND 1
REWIND 2
IF (COUNT,GT,COUNT) STOP
DO 48 (I,1,N)
COUNT=COUNT+1
DO 48 (I,1,N)
48  STORE (DATA(I))
READ (12) (DATA(I),1,N)
C      ***
C      ***
C      ***
DO 49 (I,1,N)
49  TRAINING(JUMPOVER)=DATA(I)
50  WRITE (10) (DATA(I),1,N)
C      ***
C      PLACE DATA IN ALTERNATE FROM W INPUT IF LA IS SPECIFIED ON
C      THE PHASE EXEC CARD
C      ***
51  IF (LA) AND (COUNT,GT,0) GO TO 52

```

[illegible]

36

00370 01 05-25-69 21.006

CNDRX EXECUTION REPORT

```

      K2=H                                     E 431
      JJJ=JJJ+1                               E 432
      JJJ=JJJ+K                               E 433
      DO 75 I=JJJ,1JJJ                       E 434
      WRITE (43,92) I                          E 435
      WRITE(43,93) (STAT(J),J=1,42)
09  FORMAT(131A)
      K1=K2+1                                  E 437
07  K2=K2+H                                  E 438
      NTRY=K                                    E 439
      IF (INT.QT,N) GO TO 73                   E 440
      ***                                     E 441
      C                                     E 442
      C   . . . . .                           E 443
      C   WRITE OUT TRAINING REGIONS CENTERED ON PAGE E 444
      C   ***                                     E 445
06  IF (15) GO TO 7A                           E 446
      CALL TRAME (TRAIN,M,DATA)                 E 447
      WRITE (1) (DATA(I),I=1,M)                 E 448
      J=1
      JJJ=1
      L=(130-M)/2
      IF (.NOT.111) L=(70-M)/2
      CALL MNAME
      WRITE (6,91)                               E 452
      DO 77 I=1,M2                               E 453
      WRITE (6,91) (BLANK,I=1,L),(TRAIN(I),I=J,JJ) E 455
      JJJ=1
07  JJJ=JJJ+1
      REWIND 3
07  IF (110) WRITE (43,94)
      CALL MNAME
      C                                     E 459
      C   ***                                     E 460
      C   PRINT AND PUNCH QUANTIFIED DATA IF DESIRED E 461
      C   ***                                     E 462
      WRITE (6,93)                               E 463
      IF 11 IS TRUE THEN COMPTS WAS SPECIFIED E 464
      IF (11) GO TO A0
      DO 79 I=1,1
      WRITE(6,77) I
      READ (3) (DATA(J),J=1,M)
      IF (110) WRITE (43,95) (DATA(J),J=1,M)
09  WRITE (6,92) (DATA(J),J=1,M)
      DO TO A2
00  DO 81 I=1,M
      READ (3) (DATA(J),J=1,1)
      IF (110) WRITE (43,95) (DATA(J),J=1,M)
01  WRITE (6,92) (DATA(J),J=1,M)
02  IF (.NOT.11A.AND.(110) GO TO 77
      KROMANX(KR,KK)
      REWIND 1
      CALL CMNT
03  WRITE (6,96)
      WRITE (6,96)
      WRITE (6,96)
      WRITE (6,97) KR,KR
      WRITE (6,96)
      WRITE (6,96)
      WRITE (6,96)
      STOP
      C                                     E 486
      C   ***                                     E 487
      C   ***                                     E 488
      C   ***                                     E 489
04  FORMAT (15HALTERNATE PUNCH)                E 490
05  FORMAT (17A5)                               E 491
06  FORMAT (1X,15HORIGINAL DATA/)              E 492
07  FORMAT (46,3X,10(1X,11C))                  E 493
08  FORMAT (A01)                                E 494
09  FORMAT (2HPHOTO7,24,13)                     E 495
10  FORMAT (1516)                               E 496
11  FORMAT (1X,151A1)                           E 497
12  FORMAT (10(1X,11C))                         E 498
13  FORMAT (1X,14HQUANTIFIED DATA/)            E 499
14  FORMAT (27HQUANTIFIED DATA FORMAT(6110))   E 500
15  FORMAT (6110)                               E 501
16  FORMAT (102H *****-*****-*****-***** E 502
17  *****-*****-*****-*****-*****-***** E 503
18  *****-*****-*****-*****-*****-***** E 504
19  FORMAT (44H HUF OF US CAN NOT COUNT I FIND THAT 74686 ARE 13.27M E 505
20  TRAINING REGIONS YOU FORMED,13)             E 506
21  FORMAT (//)                                  E 507
      END

```



```

1      SUBROUTINE FPM (LSL,LARGE,M,N)
2      DIMENSION SS(20), SS1(20), LSL(20), LARGE(20), IDATA(15)
3      COMMON FPM1,FPM2(15)
4      COMMON F
5      COMMON FMT1(15),FMT2(3),FMT3(5),FMT4(10),FMT5(10),FMT6(10),FMT7(10),FMT8(10),FMT9(10),FMT10(10)
6      COMMON L0,L1,L2,L3,L4,L5,L6,L7,L8,L9,L10,L11,L12,L13,L14
7      COMMON M,L,B(1000),DATA(1000)
8      EQUIVALENCE (IDATA(1),DATA(1))
9      INTEGER M
10     LOGICAL L0,L1,L2,L3,L4,L5,L6,L7,L8,L9,L10,L11,L12,L13,L14
11     REMIND 9
12     REMIND 21
13
14     C      ***
15     C      L14 IS TRUE IF ABSOLUTE QUANTIZATION
16     C      PHOTO FORM
17     C      M IS THE NUMBER OF PHOTOGRAPHS
18     C      N IS THE NUMBER OF DATA POINTS
19     C      READ IN DATA
20     C      ***
21     S=1.E20
22     B=-1.E20
23     C      GO THRU EACH PHOTOGRAPH
24     DO 3 I=1,M
25     READ (9) (IDATA(I),J=1,N)
26     C      ***
27     C      SEARCH FOR MAX AND MIN VALUES IN DATA
28     C      ***
29     DO 1 J=1,N
30     IF (S.LT.IDATA(I,J).AND.IDATA(I,J).GT.B) S=IDATA(I,J)
31     IF (B.GT.IDATA(I,J).AND.IDATA(I,J).LT.S) B=IDATA(I,J)
32     IF (L14) GO TO 3
33     B=(S+B)/2
34     C      ***
35     C      QUANTIZE AND STORE DATA ON 21
36     C      QUANTIZE RELATIVELY
37     C      ***
38     DO 2 J=1,N
39     IDATA(I,J)=(IDATA(I,J)-B)/S
40     S=1.E20
41     B=-1.E20
42     3 WRITE (21) (IDATA(I),J=1,N)
43     REMIND 9
44     REMIND 21
45     C      ***
46     C      STORE DATA BACK ON 9
47     C      FOR EACH PHOTOGRAPH
48     C      ***
49     DO 5 I=1,M
50     READ (21) (IDATA(I),J=1,N)
51     IF (L14) GO TO 5
52     C      QUANTIZE ABSOLUTELY
53     S=(S+B)/2
54     DO 4 J=1,N
55     IDATA(I,J)=(IDATA(I,J)-B)/S
56     5 WRITE (9) (IDATA(I),J=1,N)
57     DO 7 I=1,10
58     LSL(I)=0
59     LARGE(I)=10000000000
60     RETURN
61     ENTRY FPM(LSL,LARGE,M,N)
62     REMIND 9
63     REMIND 21
64     C      ***
65     C      CORPUS FORM
66     C      M IS THE NUMBER OF PHOTOGRAPHS
67     C      N IS THE NUMBER OF DATA POINTS
68     C      L14 IS TRUE FOR ABSOLUTE QUANTIZATION
69     C      INITIALIZE
70     C      ***
71     DO 8 I=1,M
72     SS(I)=1.E20
73     B=SS(I)-1.E20
74     C      ***
75     C      SEARCH FOR EXTREME VALUE
76     C      ***
77     C      GO THRU EACH POINT
78     DO 9 I=1,M
79     C      READ IN ALL THE COMPONENTS FOR THAT POINT
80     READ (9) (IDATA(I),J=1,N)
81     WRITE (21) (IDATA(I),J=1,N)
82     DO 9 J=1,N
83     IF (SS(I).GT.IDATA(I,J).AND.IDATA(I,J).GT.B) SS(I)=IDATA(I,J)
84     IF (B.GT.IDATA(I,J).AND.IDATA(I,J).LT.S) B=IDATA(I,J)
85     IF (L14) GO TO 12
86     S=1.E20
87     B=-1.E20
88     DO 10 I=1,M
89     IF (SS(I).GT.S) S=SS(I)
90     IF (B.LT.B) B=B
91     DO 11 I=1,M
92     SS(I)=S
93     11 B=SS(I)
94     12 REMIND 9
95     REMIND 21
96     C      ***
97     C      QUANTIZE
98     C      ***
99     DO 12 I=1,M
100    B=SS(I)-1.E20
101    DO 13 I=1,M
102    B=SS(I)-1.E20
103    READ (21) (IDATA(I),J=1,N)
104    C      ***
105    C      STORE DATA ON 9
106    C      ***
107    DO 14 J=1,N
108    IDATA(I,J)=(IDATA(I,J)-B)/S
109    14 WRITE (9) (IDATA(I),J=1,N)
110    DO 15 I=1,M
111    B=SS(I)-1.E20

```

\*\*\*\*\*  
 23000 WORDS OF MEMORY USED BY THIS COMPILE  
 \*\*\*\*\*

1	SUBROUTINE DEF (L,M,N,IC,M)	0	1
2	DIMENSION L(2), M(2), N(2)	0	2
3	JF=1-1000000	0	3
4	C                   ***	0	4
5	C SET SINGLE CHARACTER TRAINING REGION	0	5
6	C DEFINE IN ALPHABETICAL ORDER	0	6
7	C                   ***	0	7
8	J3=0	0	8
9	NC=0	0	9
10	DO 2 J=1,IC	0	10
11	JJ=J+J	0	11
12	DO 1 I=1,JJ	0	12
13	JF=JF+1	0	13
14	J3=J3+1	0	14
15	1 L(I)=M(JJ)	0	15
16	NC=NC+JJ	0	16
17	2 JF=JF+1-1000000	0	17
18	IF(JF)	0	18
19	END	0	19

17964 01 69-07-69 10.200

1	SUBROUTINE TRAIN (N,M,DATA)	M	1
2	COMMON MPCL,FPM7(1:12)	M	2
3	COMMON F	M	3
4	COMMON FMT1(15),FMT2(15),FMT3(15),FMT4(15),FMT5(15),FMT6(15)	M	4
5	15,FMT7(15),FMT8(15),FMT9(15),FMT10(15),FMT11(15),FMT12(15)	M	5
6	DIMENSION A(2), B(30), D(30)	M	6
7	INTEGER DATA	M	7
8	DIMENSION DATA(8)	M	8
9	INTEGER F	M	9
10	DATA D(1)=1,D(2)=1,D(3)=1,D(4)=1,D(5)=1,D(6)=1,D(7)=1,D(8)=1,D(9)=1,D(10)=1	M	10
11	D(11)=1,D(12)=1,D(13)=1,D(14)=1,D(15)=1,D(16)=1,D(17)=1,D(18)=1,D(19)=1,D(20)=1	M	11
12	D(21)=1,D(22)=1,D(23)=1,D(24)=1	M	12
13	C                   ...	M	13
14	C                   CREATE CROSS REFERENCE TABLE FOR TRAINING REGIONS	M	14
15	C                   ...	M	15
16	CALL MAKE	M	16
17	WRITE (6,FMT1) D(1),A(1)	M	17
18	B(1)=A(1)	M	18
19	A(1)=D(1)	M	19
20	DATA (1)=1	M	20
21	K=1	M	21
22	K=1	M	22
23	DO 3 1=D,M	M	23
24	IF (A(1),D(1),B(1)) GO TO 2	M	24
25	K=K+1	M	25
26	IF (M,L,B,M) GO TO 1	M	26
27	WRITE (6,FMT2) D(1),A(1)	M	27
28	B(1)=A(1)	M	28
29	A(1)=D(1)	M	29
30	DATA (1)=M	M	30
31	K=K+1	M	31
32	GO TO 3	M	32
33	2 A(1)=D(1)	M	33
34	DATA (1)=M	M	34
35	3 K=1	M	35
36	RETURN	M	36
37	END	M	37

17560 02 05-02-00 10.270

34

Line	Code	Statement	Address
1		SUBROUTINE PATTERN	J 1
2		COMMON NP(1,FORM(1:12))	J 2
3		COMMON F	J 3
4		COMMON FMT(1(15),FMT2(15),FMT3(15),FMT4(15),FMT5(15),FMT6(15),FMT7(15),FMT8(15),FMT9(15),FMT10(15),FMT11(15),FMT12(15),FMT13(15),FMT14(15),FMT15(15))	J 4
5		COMMON L(1,9),L10-L11-L12-L13-L14	J 5
6		COMMON NP,NM,LEVEL(7),NPOINT,ITAPE,AAA,L7,L1,L2,L3,L4,L5,L6,N,M1,	J 6
7		1M2,NDUMPY,L7,M5,M6,AC,ZZZ	J 7
8		***	J 8
9	C	NUMBER OF TRAINING REGIONS	J 9
10	C	LEVEL NUMBER OF LEVELS OF QUANTIZATION	J 10
11	C	NP=NUMBER OF PHOTOS	J 11
12	C	N2=NUMBER OF SIGNIFICANT FIGURES PER DATA POINT	J 12
13	C	NPOINT=ORDERED LIST OF POINTS PER TRAINING REGION	J 13
14	C	INPUT IS EXPECTED ON TAPE UNIT 'ONE'	J 14
15	C	LOGICAL TRUE FOR NORMALIZED AND FALSE FOR UNNORMALIZED	J 15
16	C	***	J 16
17	C	DIMENSION AAA(125)	J 17
18		DIMENSION ZZZ(24000), IPT(2), IREM(2), NREG(2), NPOINT(150), NPT(150)	J 18
19		15, RESULT(125), SIGMA(150), FORM(12), PLANN(1)	J 19
20		EQUIVALENTS (ZZZ(1),IPT(1), (ZZZ(1),IREM(1), (ZZZ(1),NREG(1))	J 20
21		LOGICAL L7,L1,L2,L3,L4,L5,L6,L7	J 21
22		LOGICAL L8,L9,L10,L11,L12,L13	J 22
23		INTEGER F	J 23
24		DATA PLANK/6HUNORMA,5MLIZED,6HUNORMAL,6HIZED/	J 24
25		RE=ND 1	J 25
26		RE=ND 2	J 26
27		RE=ND 3	J 27
28		RE=ND 10	J 28
29		RE=ND 20	J 29
30		***	J 30
31	C	PRINT HEADING FOR PUNCH	J 31
32	C	WHICH IS ABSOLUTE MAX NUMBER OF POINTS IN A SINGLE TRAINING REGION	J 32
33	C	***	J 33
34	C	WRITE (10,4)	J 34
35		WRITE (13,32)	J 35
36		NHIGH=0	J 36
37		NTOTAL=0	J 37
38		***	J 38
39	C	NPOINT CONTAINS NO. OF POINTS PER TRAINING REGION	J 39
40	C	COUNT TOTAL PTS. FIND MAX IN NPOINT	J 40
41	C	***	J 41
42	C	DO 1 1=1,N	J 42
43		NTOTAL=NTOTAL+NPOINT(I)	J 43
44		IF (NPOINT(I),LT,NHIGH) GO TO 1	J 44
45		NHIGH=NPOINT(I)	J 45
46		1 CONTINUE	J 46
47		J=1	J 47
48		IF (J) J=3	J 48
49		N=J-1	J 49
50		CALL NAME	J 50
51		WRITE (6,33) (PLAN(I),I=J,N)	J 51
52		WRITE (6,34)	J 52
53		***	J 53
54	C	PRINT PARAMETERS	J 54
55	C	***	J 55
56	C	WRITE (6,35) N,LEVEL,NP	J 56
57	C	***	J 57
58	C	INITIALIZE	J 58
59	C	***	J 59
60	C	DO 2 1=1,100	J 60
61		2 RESULT(1)=0.	J 61
62		***	J 62
63	C	RESULT IS PROBABILITY MATRIX	J 63
64	C	CALCULATE NO. OF LOCATIONS FOR USE IN IREM AND IPT	J 64
65	C	NEVER MORE THAN 1 IN BOTTOM OF ZZZ	J 65
66	C	***	J 66
67	C	N=20000-NHIGH	J 67
68		IF (N,LT,10000) N=10000	J 68
69		N=0	J 69
70		N=0	J 70
71		IP=0	J 71
72		IP=0	J 72
73		IP=0	J 73
74	C	***	J 74
75	C	CALCULATE NORMALIZING FACTORS	J 75
76	C	***	J 76
77		DO 3 1=1,N	J 77
78		SIGMA(1)=1/FLOAT(NPOINT(I))	J 78
79		3 IP=IP+LOCAT(NPOINT(I))	J 79
80		***	J 80
81	C	POINT IS A FUNCTION OF THE POSITION IN THE TOP PART OF ZZZ	J 81
82	C	INDICATION WHERE THE PRESENT UNIQUE N-TUPLE IS UNIQUE TO YOU	J 82
83	C	WILL BE STORED	J 83
84	C	THIS IS THE DESIGNATION OF THE BLOCK #1 IN THE TOP HALF OF ZZZ	J 84
85	C	IP=1 IS NUMBER OF POINTS IN A1 TRAINING REGION	J 85
86	C	AND LOAD IT UP IN THE BOTTOM PART OF ZZZ	J 86
87	C	STARTING AT THE TOP AND COMING DOWN	J 87
88	C	***	J 88
89		ACCOUNT=1	J 89
90		LICH=1	J 90
91	C	***	J 91
92	C	41 RANGES OVER NO. OF TRAINING REGIONS	J 92
93	C	***	J 93
94		DO 4 1=1,N	J 94
95		1POINT(NPOINT(I))	J 95
96	C	***	J 96
97	C	READ IN SINGLE TRAINING REGION	J 97
98	C	***	J 98
99	C	READ IN (NREG(I),1,1,IP(I))	J 99
100	C	***	J 100
101	C	SET N-TUPLE IN GIVEN LOCATIONS	J 101
102	C	PUT NO. TIMES OCCURRED IN 2ND LOCATIONS	J 102
103	C	TRAINING REGION IS STORED IN LOWER HALF OF 2ND LOCATIONS	J 103

17464 03 09-02-60 10.284

```

104 C *** J 104
105 DO 7 L=1,1001 J 105 01
106 ***** J 106 02
107 IF (L=10,HCOUNT) GO TO 9 J 107 03
108 HCOUNT=1 J 108 04
109 DO 4 I=1,4,11 J 109 05
110 I=20002-I+2 J 110 06
111 IF (HRC,13,IREM(I)) GO TO 6 J 111 07
112 CONTINUE J 112 08
113 C *** J 113 09
114 IF OVERLAP OCCURS IN SEARCH AREA GO TO 777 AND STORE RESULTS J 114 10
115 C HRC(I) HAS NOT OCCURRED BEFORE IN THE TRAINING REGION, NAME J 115 11
116 POSITION TO STORE IT J 116 12
117 C *** J 117 13
118 9 IF (HCOUNT=2,ST,05) GO TO 97 J 118 14
119 HCOUNT=20002-HCOUNT+2 J 119 15
120 IREM(HCOUNT)=HRC J 120 16
121 IPT(HCOUNT-1)=10000+4*I+100 J 121 17
122 HCOUNT=HCOUNT+1 J 122 18
123 GO TO 7 J 123 19
124 C *** J 124 20
125 HRC(I) HAS OCCURRED BEFORE IN TO H1 J 125 21
126 C *** J 126 22
127 6 I=20001-I+2 J 127 23
128 IPT(I)=IPT(I)+10000 J 128 24
129 CONTINUE J 129 25
130 LICK=HCOUNT J 130 26
131 H=HCOUNT+1 J 131 27
132 C *** J 132 28
133 H2 IS THE LOWEST LOCATION ON THE TOP PART WHICH IS USED J 133 29
134 C *** J 134 30
135 I=20000 J 135 31
136 H2=20001-I J 136 32
137 C *** J 137 33
138 STORE RESULTS ON TAPE 10 J 138 34
139 C *** J 139 35
140 WRITE (10) I J 140 36
141 WRITE (10) (IREM(I),I=H2,20000) J 141 37
142 IF (HRC,10,0) GO TO 11 J 142 38
143 JTAPE=2 J 143 39
144 LTAPE=3 J 144 40
145 RE=100.2 J 145 41
146 RE=100.0 J 146 42
147 LL=1 J 147 43
148 10 READ (JTAPE) HCOUNT J 148 44
149 H=20001-HCOUNT J 149 45
150 HCOUNT=1 J 150 46
151 HRC=2-H J 151 47
152 READ (JTAPE) (IREM(I),I=H,100) J 152 48
153 HRC=1 J 153 49
154 11 LEND=1 J 154 50
155 C *** J 155 51
156 C SEQUENCE N-TUPLES IN DECREASING ORDER J 156 52
157 C *** J 157 53
158 DO 12 I=1,4 J 158 54
159 IPT(I)=1 J 159 55
160 DO 12 J=1,100 J 160 56
161 I=21002-I+2 J 161 57
162 J=20002-J+2 J 162 58
163 IF (IREM(I),10,IREM(J)) GO TO 13 J 163 59
164 IPT(I)=IPT(I)+1 J 164 60
165 IREM(I)=IREM(J) J 165 61
166 IREM(J)=IPT(I) J 166 62
167 IPT(I)=IPT(I)-1 J 167 63
168 IPT(J)=IPT(J)+1 J 168 64
169 IPT(J)=IPT(J)-1 J 169 65
170 CONTINUE J 170 66
171 IF (HRC,10,0) GO TO 15 J 171 67
172 WRITE (JTAPE) (IREM(I),I=H,100) J 172 68
173 LL=LL+1 J 173 69
174 IF (LL,10,HRC) GO TO 16 J 174 70
175 HPT=H J 175 71
176 HRC=IREM(H) J 176 72
177 DO 13 I=1,10000,2 J 177 73
178 IF (IREM(I),H,100) GO TO 14 J 178 74
179 CONTINUE J 179 75
180 14 H=H/2 J 180 76
181 CONTINUE J 181 77
182 HCOUNT=0 J 182 78
183 15 J 183 79
184 C *** J 184 80
185 C LOOK FOR STARTING AND END POINTS OF EACH N-TUPLE J 185 81
186 C LOOK FROM PRESENT N-TUPLE DOWN J 186 82
187 C *** J 187 83
188 DO 17 I=1,100 J 188 84
189 J=1 J 189 85
190 I=20002-I+2 J 190 86
191 J=20002-J+2 J 191 87
192 IF (IREM(I),H,IREM(J)) GO TO 18 J 192 88
193 CONTINUE J 193 89
194 J=J+1 J 194 90
195 17 J=J-1 J 195 91
196 C *** J 196 92
197 C J IS A FUNCTION OF THE LAST N-TUPLE LIES TWO 1 TO N-TUPLE FOR ONE J 197 93
198 C *** J 198 94
199 18 1 J 199 95
200 IF (I,10,0) GO TO 19 J 200 96
201 C *** J 201 97
202 N-TUPLE OCCURS IF HRC THEN ONE TO J 202 98
203 C *** J 203 99
204 19 20 J 204 100
205 C *** J 205 101
206 IF ONE OR TWO, IN N-TUPLE CLASSIFY N-TUPLE J 206 102
207 C *** J 207 103

```

200		10 15-20000-100	J 200 177
201	C	"	J 201
202	C	15 IS THE LOCATION OF THE COUNT FOR THE 1 TO N-TUPLE	J 202
203	C	"	J 203
204	C	WRITE(15,10000)	J 204 178
205	C	WRITE(15,1010000/100)	J 205 179
206	C	"	J 206
207	C	N IS NUMBER OF TIMES THAT THE N-TUPLE IS IN OCCURED	J 207
208	C	"	J 208
209	C	GO TO 21	J 209 180
210	C	"	J 210
211	C	IF TWO OR MORE T.O. IN N-TUPLE LEAD FOR PROBABILITY T.O.	J 211
212	C	AND CLASSIFY	J 212
213	C	"	J 213
214	C	20 CALL FUNCTION F(J,I,P,T,STIME,N,4,0,0,0,1,TOTAL,SPROB)	J 214 181
215	C	20 20 20 20 20	J 215 182
216	C	WRITE(16,10000)	J 216 183
217	C	WRITE(16,1010000/100)	J 217 184
218	C	"	J 218
219	C	STORE IN WORD N-TUPLE AND CLASSIFICATION	J 219
220	C	IF WORD FULL GO TO 201 AND STORE WORD	J 220
221	C	"	J 221
222	C	20 IF (COUNT(OT,41000)) GO TO 20	J 222 187
223	C	15-20000-100	J 223 188
224	C	WRITE(16,10000)	J 224 189
225	C	WRITE(16,1010000/100)	J 225 190
226	C	"	J 226
227	C	IF (1,1,1,1,1) GO TO 10	J 227 191
228	C	COUNT=COUNT+1	J 228 192
229	C	IF (1,1,1,1,1) GO TO 20	J 229 193
230	C	"	J 230
231	C	CLASSIFY ALL N-TUPLES	J 231
232	C	GO TO 10 FOR THE N-TUPLE	J 232
233	C	HOW IS THE CLASSIFICATION BY BAYES' RULE	J 233
234	C	IS IN POSITION IN RESULT ADDR FOR (0,0,0,0)	J 234
235	C	"	J 235
236	C	GO TO 10,100	J 236 194
237	C	15-20000-100	J 237 195
238	C	WRITE(16,10000)	J 238 196
239	C	WRITE(16,1010000/100)	J 239 197
240	C	"	J 240
241	C	STORE IN WORD N-TUPLE AND CLASSIFICATION	J 241
242	C	IF WORD FULL GO TO 201 AND STORE WORD	J 242
243	C	"	J 243
244	C	20 IF (COUNT(OT,41000)) GO TO 20	J 244 197
245	C	15-20000-100	J 245 198
246	C	WRITE(16,10000)	J 246 199
247	C	WRITE(16,1010000/100)	J 247 200
248	C	"	J 248
249	C	STORE IN WORD N-TUPLE AND CLASSIFICATION	J 249
250	C	IF WORD FULL GO TO 201 AND STORE WORD	J 250
251	C	"	J 251
252	C	20 IF (COUNT(OT,41000)) GO TO 20	J 252 203
253	C	15-20000-100	J 253 204
254	C	WRITE(16,10000)	J 254 205
255	C	WRITE(16,1010000/100)	J 255 206
256	C	"	J 256
257	C	STORE IN WORD N-TUPLE AND CLASSIFICATION	J 257
258	C	IF WORD FULL GO TO 201 AND STORE WORD	J 258
259	C	"	J 259
260	C	20 IF (COUNT(OT,41000)) GO TO 20	J 260 207
261	C	15-20000-100	J 261 208
262	C	WRITE(16,10000)	J 262 209
263	C	WRITE(16,1010000/100)	J 263 210
264	C	"	J 264
265	C	STORE IN WORD N-TUPLE AND CLASSIFICATION	J 265
266	C	IF WORD FULL GO TO 201 AND STORE WORD	J 266
267	C	"	J 267
268	C	20 IF (COUNT(OT,41000)) GO TO 20	J 268 211
269	C	15-20000-100	J 269 212
270	C	WRITE(16,10000)	J 270 213
271	C	WRITE(16,1010000/100)	J 271 214
272	C	"	J 272
273	C	STORE IN WORD N-TUPLE AND CLASSIFICATION	J 273
274	C	IF WORD FULL GO TO 201 AND STORE WORD	J 274
275	C	"	J 275
276	C	20 IF (COUNT(OT,41000)) GO TO 20	J 276 215
277	C	15-20000-100	J 277 216
278	C	WRITE(16,10000)	J 278 217
279	C	WRITE(16,1010000/100)	J 279 218
280	C	"	J 280
281	C	STORE IN WORD N-TUPLE AND CLASSIFICATION	J 281
282	C	IF WORD FULL GO TO 201 AND STORE WORD	J 282
283	C	"	J 283
284	C	20 IF (COUNT(OT,41000)) GO TO 20	J 284 219
285	C	15-20000-100	J 285 220
286	C	WRITE(16,10000)	J 286 221
287	C	WRITE(16,1010000/100)	J 287 222
288	C	"	J 288
289	C	STORE IN WORD N-TUPLE AND CLASSIFICATION	J 289
290	C	IF WORD FULL GO TO 201 AND STORE WORD	J 290
291	C	"	J 291
292	C	20 IF (COUNT(OT,41000)) GO TO 20	J 292 223
293			

1. 2019. 10. 10. 10:00 ~ 10:30 (10:00 ~ 10:30)

1	SUBROUTINE OUTP (RESULT,N)			1
2	DIMENSION M(50*(N+1), 17*(25))			2
3	COMMON M(51),FORMT(112)			3
4	COMMON F			4
5	COMMON FMT1(19),FMT2(3),FMT9(9),FMT11(4),FMT12(7),FMT13(3),FMT14(2			5
6	1),FMT1A(2),FMT22(3),FMT5A(4),FMT57(4),FMT5B(4)			6
7	COMMON LA,L9,L10,L11,L12,L13,L14			7
8	INTEGER F			8
9	DATA 17,FMT1,FMT2,FMT9,FMT11,FMT12,FMT13,FMT14,FMT1A,FMT22,FMT5A,FMT57,FMT5B,LA,L9,L10,L11,L12,L13,L14			9
10	F			10
11	C			11
12	C			12
13	C			13
14	11=1			14
15	111=0			15
16	NN=0			16
17	1 CALL NAME			17
18	NN=0			18
19	IF (KULT,A) GO TO 9			19
20	LINE=0			20
21	RA=0			21
22	NS=25-N			22
23	IF (NS.LD.0) GO TO 3			23
24	DO 2 1=1,NS			24
25	LINE=LINE+1			25
26	2 WRITE (6,19)			26
27	3 LINE=LINE+7			27
28	WRITE (6,FMT5A)			28
29	WRITE (6,FMT12)			29
30	WRITE (6,FMT13) (1,1011,111)			30
31	IF (L11R,NF,24) GO TO 4			31
32	WRITE (6,FMT14) 10L(1)			32
33	4 DO 6 J=1,N			33
34	LINE=LINE+1			34
35	WRITE (6,FMT57) J,(RESULT(1,J),1011,1			35
36	IF (44-LINE,L8,0) GO TO 5			36
37	L9=LINE-23			37
38	IF (L11L,0) GO TO 9			38
39	WRITE (6,FMT14) 10L(L)			39
40	4 LINE=LINE+1			40
41	WRITE (6,19)			41
42	IF (44-LINE,L7,0) GO TO 6			42
43	L9=LINE-23			43
44	IF (L11L,0) GO TO 6			44
45	WRITE (6,FMT14) 10L(L)			45
46	6 CONTINUE			46
47	7 LINE=LINE+1			47
48	L9=LINE-23			48
49	IF (L11L,23) GO TO 8			49
50	WRITE (6,FMT16) 10L(L)			50
51	GO TO 7			51
52	8 NN=NN+1			52
53	11=11+8			53
54	111=111+8			54
55	IF (NS.GT.0) GO TO 1			55
56	RETURN			56
57	9 DO 10 1=1,23			57
58	10 WRITE (6,19)			58
59	WRITE (6,FMT5A)			59
60	WRITE (6,FMT12)			60
61	111=111+1			61
62	WRITE (6,FMT13) (1,1011,111)			62
63	WRITE (6,FMT16) 10L(1)			63
64	L11=24			64
65	DO 12 J=1,N			65
66	LINE=LINE+1			66
67	WRITE (6,FMT57) J,(RESULT(1,J),1011,111)			67
68	IF (44-LINE,L8,0) GO TO 11			68
69	L9=LINE-23			69
70	IF (L11L,0) GO TO 11			70
71	WRITE (6,FMT14) 10L(L)			71
72	11 LINE=LINE+1			72
73	WRITE (6,19)			73
74	IF (44-LINE,L8,0) GO TO 12			74
75	L9=LINE-23			75
76	IF (L11L,0) GO TO 12			76
77	WRITE (6,FMT14) 10L(L)			77
78	12 CONTINUE			78
79	13 LINE=LINE+1			79
80	L9=LINE-23			80
81	IF (L11L,23) GO TO 14			81
82	WRITE (6,FMT16) 10L(L)			82
83	GO TO 13			83
84	14 RETURN			84
85	C			85
86	C			86
87	C			87
88	19 FORMAT (2)			88
89	6A			89
90				90

\*\*\*\*\*  
 = Page 4023 of 4029, 138 of 140 competition  
 \*\*\*\*\*

17564 03 05-02-49 1A.114

1	SUBROUTINE DECISION (I,J,M1,SIGMA,H,N,A,LT,NTOTAL,VRPHGH)		
2	DIMENSION M1(2), SIGMA(2)	L	1
3	DIMENSION A(2), C(38)	L	2
4	LOGICAL LT	L	3
5	M=N	L	4
6	DO 1 K=1,N	L	5
7	1 C(K)=0.	L	6
8	VRPHGH=0.	L	7
9	ZLARGE=.1E+21	L	8
10	DO 2 K=1,J	L	9
11	K=20001-2*K	L	10
12	LL=M1(K)/10000	L	11
13	LL=M1(K)-LL*10000/100	L	12
14	2 C(L)=C(L)+FLOAT(LL)	L	13
15	VRPHGH=FLOAT(M1)/FLOAT(NTOTAL)	L	14
16	M=1	L	15
17	C	L	16
18	C	L	17
19	C	L	18
20	IF NORMALIZED IS SPECIFIED MULT BY NORMALIZING FACTORS	L	19
21	IF (.NOT. LT) GO TO 4	L	20
22	DO 3 L=1,N	L	21
23	IF (C(L).LT.1.E-19) C(L)=0.	L	22
24	3 C(L)=C(L)*SIGMA(L)	L	23
25	4 DO 6 L=1,N	L	24
26	SMALL=0.	L	25
27	C	L	26
28	C	L	27
29	C	L	28
30	DO 5 K=1,N	L	29
31	SMALL=SMALL+AMIN0(C(K))	L	30
32	5 N=K+1	L	31
33	SMALL=SMALL-10000.	L	32
34	C	L	33
35	C	L	34
36	C	L	35
37	LOOK FOR PNEUMONIT TRAINING REGION	L	36
38	IF (SMALL.GT.ZLARGE)	L	37
39	ZLARGE=SMALL	L	38
40	NEL	L	39
41	6 CONTINUE	L	40
42	C	L	41
43	C	L	42
44	C	L	43
	N IS THE RESULTING TRAINING REGION	L	44
	RETURN	L	45
	END	L	46

45





47

```

S   FORTRAN NDECK.COMDK                               SEARCH
    SUBROUTINE SEARCH (KPT,KLAS,KK,SYMBOL,J2,NP,K,LEVEL)
    DIMENSION KPT(2), KLAS(2), SY(16)                P 2
    DIMENSION KP(100)                                  P 4
    DIMENSION LIST(2)                                  P 5
    COMMON /IST/ IST                                    P 6
    DATA SY/1HA,1HP,1HC,1HD,1HE,1HF,1HG,1HH,1HI,1HJ,1HK,1HL,1HM,1HN,1HO,1HP,1HQ,1HR,1HS,1HT,1HU,1HV,1HW,1HX,1HY,1HZ,1MI,1M2,1M3,1M4,1M5,
    21MA,1M7,1MC,1M9,1MO/                               P 7
C   K IS THE CATEGORY NUMBER                            P 8
C   SYMBOL IS THE CATEGORY CODE
C   KK IS THE COMPACTED NTUPLE
C   J2 IS UPPER LIMIT FOR KPT AND KLAS ARRAYS
C   LOOK FOR COMPACTED N-TUPLE KK IN LIST
    KS=0                                                P 9
    KL=J2+1                                             P 10
    KTRY=(KS+KL)/2                                       P 11
1  CONTINUE
    IF (KK-KPT(KTRY)) 3,4,2
2  IF ((KL-KS).LE.1) GO TO 5
    KL=KTRY
    KTRY=(KL+KS+1)/2
    GO TO 1
3  IF (KL-KS.LE.1) GO TO 5
    KS=KTRY
    KTRY=(KL+KS)/2
    GO TO 1
4  K=KLAS(KTRY)
    SYMBOL=SY(K)
    RETURN                                             P 13
5  CONTINUE                                             P 14
C   KK CANNOT BE FOUND DO NEAREST NEIGHBOR SEARCH      P 15
    I1=IDIST(KK,KPT(1),NP,LEVEL)                       P 16
    L=1
    KP(1)=1
    DO 7 J=2,NP
        I2=IDIST(KK,KPT(J),NP,LEVEL)                   P 17
        IF (I1.LT.I2) GO TO 7
        IF (I1.EQ.I2) GO TO 6
        IF (L.GT.99) GO TO 7
        L=L+1
        KP(L)=J
        GO TO 7
    6 CONTINUE
        I1=I2
        L=L+1
        KP(L)=J
    7 CONTINUE
    X=RCM(IST)
    K=X*FLOAT(L-1)+1.5
C   KP IS THE ARRAY CONTAINING THE INDEXES FOR ALL THOSE POINTS IN
C   THE KPT ARRAY WHICH ARE CLOSEST TO KK
C   L IS THE UPPER LIMIT OF KP
C   WE WILL CHOOSE ONE POINT FROM THE KP ARRAY AT RANDOM AND
C   IDENTIFY KK WITH THE CATEGORY OF THE POINT IN KPT ASSOCIATED WITH
C   THE RANDOMLY CHOSEN ONE
    K=KP(K)
    K=KLAS(K)
    SYMBOL=SY(K)
    RETURN
END

```

```

S   FORTRAN NDECK.COMDK                               IDIST
    FUNCTION IDIST(KG,KP,NP,LEVEL)
    IDIST=0
    KK=KG
    LL=KP
    NP1=NP-1
    DO 1 J=1,NP1
        K=KP/LEVEL
        L1=KK-K*LEVEL
        I=LL/LEVEL
        L2=LL-I*LEVEL
        KK=K
        LL=I
1  IDIST=IDIST*(L1-L2)*(L1-L2)
    IDIST=IDIST*(K-1)*(K-1)
    RETURN
END

```

28 480  
FORMAT FOR TRAINING DATA IS (20A1)

INPUT FORMAT FOR DATA WAS (2011)  
OUTPUT FORMAT FOR DATA IS (1X,2011)

[illegible]





### TEST DATA

## TABLE TABLE

## PLOT SYMBOLS VS INPUT SYMBOLS

A

**B**

### TEST DATA

## TRAINING REGIONS

[illegible][illegible]

TEST DATA

NO. TRAINING REGIONS	2
NORMALIZED NO. LEVELS	4
NO. PHOTOS	2

TEST DATA

THE PROBABILITY MATRIX

TRAINING REGION CLASSIFIED AS

	1	2
1	0.4545E 00	0.5455E 00
2	0.1111E 00	0.8889E 00

TOTAL PROBABILITY

0.0500E 00

TRAINING REGION

N-TUPLES

35	1	2
28	0	400
21	400	400
14	400	400
7	800	200
35	400	400
0	2	0
0	28	28
0	1	1



PHOTO CLASSIFICATION

[illegible]

### CONTINGENCY TABLE

TRAINING REGION CLASSIFIED AS

		1	2
T			
0	1	0.10000 04	0.12000 04
U			
0	2	0.20000 03	0.14000 04

**YAKUTIA**

SECOND HALF, NUMBER OF POINTS IS

•••••

[illegible][illegible]

Page	1	2	3	4	5	6	7	8	9
1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1	1
5	1	1	1	1	1	1	1	1	1
6	1	1	1	1	1	1	1	1	1
7	1	1	1	1	1	1	1	1	1
8	1	1	1	1	1	1	1	1	1
9	1	1	1	1	1	1	1	1	1
10	1	1	1	1	1	1	1	1	1
11	1	1	1	1	1	1	1	1	1
12	1	1	1	1	1	1	1	1	1
13	1	1	1	1	1	1	1	1	1
14	1	1	1	1	1	1	1	1	1
15	1	1	1	1	1	1	1	1	1
16	1	1	1	1	1	1	1	1	1
17	1	1	1	1	1	1	1	1	1
18	1	1	1	1	1	1	1	1	1
19	1	1	1	1	1	1	1	1	1
20	1	1	1	1	1	1	1	1	1
21	1	1	1	1	1	1	1	1	1
22	1	1	1	1	1	1	1	1	1
23	1	1	1	1	1	1	1	1	1
24	1	1	1	1	1	1	1	1	1
25	1	1	1	1	1	1	1	1	1
26	1	1	1	1	1	1	1	1	1
27	1	1	1	1	1	1	1	1	1
28	1	1	1	1	1	1	1	1	1
29	1	1	1	1	1	1	1	1	1
30	1	1	1	1	1	1	1	1	1
31	1	1	1	1	1	1	1	1	1
32	1	1	1	1	1	1	1	1	1
33	1	1	1	1	1	1	1	1	1
34	1	1	1	1	1	1	1	1	1
35	1	1	1	1	1	1	1	1	1
36	1	1	1	1	1	1	1	1	1
37	1	1	1	1	1	1	1	1	1
38	1	1	1	1	1	1	1	1	1
39	1	1	1	1	1	1	1	1	1
40	1	1	1	1	1	1	1	1	1
41	1	1	1	1	1	1	1	1	1
42	1	1	1	1	1	1	1	1	1
43	1	1	1	1	1	1	1	1	1
44	1	1	1	1	1	1	1	1	1
45	1	1	1	1	1	1	1	1	1
46	1	1	1	1	1	1	1	1	1
47	1	1	1	1	1	1	1	1	1
48	1	1	1	1	1	1	1	1	1
49	1	1	1	1	1	1	1	1	1
50	1	1	1	1	1	1	1	1	1
51	1	1	1	1	1	1	1	1	1
52	1	1	1	1	1	1	1	1	1
53	1	1	1	1	1	1	1	1	1
54	1	1	1	1	1	1	1	1	1
55	1	1	1	1	1	1	1	1	1
56	1	1	1	1	1	1	1	1	1
57	1	1	1	1	1	1	1	1	1
58	1	1	1	1	1	1	1	1	1
59	1	1	1	1	1	1	1	1	1
60	1	1	1	1	1	1	1	1	1
61	1	1	1	1	1	1	1	1	1
62	1	1	1	1	1	1	1	1	1
63	1	1	1	1	1	1	1	1	1
64	1	1	1	1	1	1	1	1	1
65	1	1	1	1	1	1	1	1	1
66	1	1	1	1	1	1	1	1	1
67	1	1	1	1	1	1	1	1	1
68	1	1	1	1	1	1	1	1	1
69	1	1	1	1	1	1	1	1	1
70	1	1	1	1	1	1	1	1	1
71	1	1	1	1	1	1	1	1	1
72	1	1	1	1	1	1	1	1	1
73	1	1	1	1	1	1	1	1	1
74	1	1	1	1	1	1	1	1	1
75	1	1	1	1	1	1	1	1	1
76	1	1	1	1	1	1	1	1	1
77	1	1	1	1	1	1	1	1	1
78	1	1	1	1	1	1	1	1	1
79	1	1	1	1	1	1	1	1	1
80	1	1	1	1	1	1	1	1	1
81	1	1	1	1	1	1	1	1	1
82	1	1	1	1	1	1	1	1	1
83	1	1	1	1	1	1	1	1	1
84	1	1	1	1	1	1	1	1	1
85	1	1	1	1	1	1	1	1	1
86	1	1	1	1	1	1	1	1	1
87	1	1	1	1	1	1	1	1	1
88	1	1	1	1	1	1	1	1	1
89	1	1	1	1	1	1	1	1	1
90	1	1	1	1	1	1	1	1	1
91	1	1	1	1	1	1	1	1	1
92	1	1	1	1	1	1	1	1	1
93	1	1	1	1	1	1	1	1	1
94	1	1	1	1	1	1	1	1	1
95	1	1	1	1	1	1	1	1	1
96	1	1	1	1	1	1	1	1	1
97	1	1	1	1	1	1	1	1	1
98	1	1	1	1	1	1	1	1	1
99	1	1	1	1	1	1	1	1	1
100	1	1	1	1	1	1	1	1	1



### TEST DATA

## TABLE TABLE

## PLOT SYMBOLS VS INPUT SYMBOLS

A

A

**R**

9

### TEST DATA

## TRAINING REGIONS

[illegible][illegible]

[illegible]

## CONTINGENCY TABLES

TRAINING REGION CLASSIFIED AS

		1	2
T R U E	1	0.1400E 04	0.1000E 04
	2	0.4000E 03	0.1200E 04
T R A I N I N G			
R E G I O N			

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D		
(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)		
1. ORIGINATING ACTIVITY (Corporate author)  Center for Research, Inc. The University of Kansas		2a. REPORT SECURITY CLASSIFICATION  Unclassified
		2b. GROUP
3. REPORT TITLE  The Bayesian Approach to Identification of a Remotely Sensed Environment		
4. DESCRIPTIVE NOTES (Type of report and inclusive dates)		
5. AUTHOR(S) (First name, middle initial, last name)  Robert M. Haralick		
6. REPORT DATE  July 1969	7a. TOTAL NO. OF PAGES  68	7b. NO. OF REFS
8a. CONTRACT OR GRANT NO.  DAAK02-68-C-0089	9a. ORIGINATOR'S REPORT NUMBER(S)  133-9	
b. PROJECT NO. ARPA Order No. 1079		
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) Supported	
d.		
10. DISTRIBUTION STATEMENT  No Limitation		
11. SUPPLEMENTARY NOTES  Funded by DOD Project Themis under ARPA Order No. 1079		12. SPONSORING MILITARY ACTIVITY Monitoring: U.S. Army Engineer Topographic Labs Geographic Information Systems Branch Geographic System Division Ft. Belvoir, Virginia
13. ABSTRACT  The first part of this paper provides a brief tutorial introduction of the Bayesian Approach to identification of a remotely sensed environment. The second part describes the input data deck setup for the Fortran IV program which has been written to implement this approach. The third part describes file usage and subroutine organization. The fourth part provides a listing of the program with a simple sample data set.		

DD FORM 1473

REPLACES DD FORM 1473, 1 JAN 64, WHICH IS OBSOLETE FOR ARMY USE.

Unclassified

Security Classification

**CRES LABORATORIES**

**Chemical Engineering Low Temperature Laboratory**

**Remote Sensing Laboratory**

**Electronics Research Laboratory**

**Chemical Engineering Heat Transfer Laboratory**

**Nuclear Engineering Laboratory**

**Environmental Health Engineering Laboratory**

**Digital Computer Technology Laboratory**

**Water Resources Institute**